Nonlinear Differ. Equ. Appl. (2023) 30:7

1021-9722/23/010001-27
*published online* November 2, 2022
https://doi.org/10.1007/s00030-022-00810-2

**Nonlinear Differential Equations and Applications NoDEA**

# Gradient formula for transition semigroup corresponding to stochastic equation driven by a system of independent Lévy processes

Alexei M. Kulik, Szymon Peszat and Enrico Priola

**Abstract.** Let $(P_t)$ be the transition semigroup of the Markov family $(X^x(t))$ defined by SDE

$$\mathrm{d}X = b(X)\mathrm{d}t + \mathrm{d}Z, \qquad X(0) = x,$$

where $Z = (Z_1, \ldots, Z_d)^*$ is a system of independent real-valued Lévy processes. Using the Malliavin calculus we establish the following gradient formula

$$\nabla P_t f(x) = \mathbb{E}\, f\left(X^x(t)\right) Y(t,x), \qquad f \in B_b(\mathbb{R}^d),$$

where the random field $Y$ does not depend on $f$. Moreover, in the important cylindrical $\alpha$-stable case $\alpha \in (0,2)$, where $Z_1, \ldots, Z_d$ are $\alpha$-stable processes, we are able to prove sharp $L^1$-estimates for $Y(t,x)$. Uniform estimates on $\nabla P_t f(x)$ are also given.

**Mathematics Subject Classification.** 60H10, 60H07, 60G51, 60J75.

**Keywords.** Bismut–Elworthy–Li formula, Lévy processes, Malliavin calculus.

## 1. Introduction

Let $(P_t)$ be the transition semigroup of a Markov family $X = (X^x(t))$ on $\mathbb{R}^d$, that is

$$P_t f(x) = \mathbb{E}\, f(X^x(t)), \qquad f \in B_b(\mathbb{R}^d),\ t \geq 0, x \in \mathbb{R}^d. \tag{1}$$

The work of Alexei Kulik was supported by Polish National Science Center Grant 2019/33/B/ST1/02923. The work of Szymon Peszat was supported by Polish National Science Center Grant 2017/25/B/ST1/02584. The work of Enrico Priola was supported by the Grant 346300 for IMPAN from the Simons Foundation and the matching 2015–2019 Polish MNiSW fund.

In the paper $X$ is given by the stochastic differential equation

$$\mathrm{d}X^x(t) = b(X^x(t))\mathrm{d}t + \mathrm{d}Z(t), \qquad X^x(0) = x \in \mathbb{R}^d, \tag{2}$$

where $b\colon \mathbb{R}^d \mapsto \mathbb{R}^d$ is a $C^2(\mathbb{R}^d, \mathbb{R}^d)$ and Lipschitz mapping and

$$Z(t) = (Z_1(t), \dots, Z_d(t))^*, \qquad t \geq 0,$$

is a Lévy process in $\mathbb{R}^d$. We assume that $Z_j$, $j = 1, \dots, d$, are independent real-valued Lévy processes. We denote by $m_j$ the Lévy measure of $Z_j$. Recall that

$$\int_{\mathbb{R}} (\xi^2 \wedge 1)\, m_j(\mathrm{d}\xi) < +\infty.$$

We assume that each $Z_j$ is of purely jump type

$$Z_j(t) = \int_0^t \int_{\{\xi\in\mathbb{R}:\ |\xi|>1\}} \xi\, \Pi_j(\mathrm{d}s, \mathrm{d}\xi) + \int_0^t \int_{\{\xi\in\mathbb{R}:\ |\xi|\leq 1\}} \xi \left[\Pi_j(\mathrm{d}s, \mathrm{d}\xi) - \mathrm{d}s m_j(\mathrm{d}\xi)\right], \tag{3}$$

where $\Pi_j(\mathrm{d}s, \mathrm{d}\xi)$ is a Poisson random measure on $[0, +\infty) \times \mathbb{R}$ with intensity measure $\mathrm{d}s m_j(\mathrm{d}\xi)$.

The main aim of this article is to establish the following gradient formula

$$\nabla P_t f(x) = \mathbb{E}\, f\left(X^x(t)\right) Y(t, x), \qquad f \in B_b(\mathbb{R}^d), \tag{4}$$

where the random field $Y$ does not depend on $f$. The gradient formulae of such type date back to [5,10] and are frequently called the *Bismut–Elworthy–Li formulae*. Note that [5] uses an approach based on the Girsanov transformation. On the other hand [10] introduces martingale methods to derive formulae like (4) in the Gaussian setting; this approach also works for jump diffusions with a non-degenerate Gaussian component (cf. Section 5 in [20]).

One important consequence of (4) is the strong Feller property of the semigroup $(P_t)$, e.g. [7–9,18], which in particular motivates our interest in this topic. Moreover, such gradient formulae allow the Greeks computations for pay-off functions in mathematical finance, e.g. [6,11]. In particular in [11] the authors apply the Malliavin calculus on the Wiener space to the sensitivity analysis for asset price dynamics models.

For Lévy-driven SDEs with a possibly degenerate Gaussian component, the Bismut–Elworthy–Li formula has been obtained in [21] under the assumption on the Lévy measure to have a density with respect to Lebesgue measure in $\mathbb{R}^d$; see also [22,23] for the Bismut–Elworthy–Li formula for an SDE driven by a subordinated Brownian motion. In our study, we are focused on the more difficult situation, where the noise is presented by a collection of one-dimensional Lévy processes, and thus is quite singular.

In plain words, the substantial complication of the problem in our case is that the class of the random vector fields, which are "admissible" for the noise in the sense that they allow the integration-by-parts formula, is much more restricted. Namely, in our case only the "coordinate axis differentiability directions" in $\mathbb{R}^d$ are actually allowed, while in the case of the Lévy measure with a density there are no limitation on these directions. For the first advances in the Malliavin calculus for Lévy noises, supported by (singular) collection of curves, we refer to [16].

In the important cylindrical $\alpha$-stable case (i.e., when each $Z_j$ is $\alpha$-stable) with $\alpha \in (0,2)$ we obtain the sharp estimate

$$\sup_{x\in\mathbb{R}^d} \mathbb{E}\,|Y(t,x)| \le C_T\, t^{-\frac{1}{\alpha}}, \qquad t \in (0,T]. \tag{5}$$

The method we use to obtain (5) seems to be of independent interest. It has two main steps. The first one is a bound for $\mathbb{E}\,|Y(t) - Y(t,x)|$, where $Y(t)$ corresponds to $Y(t,x)$ when $b = 0$ in (2), i.e., $X^x(t) = x + Z(t)$. The second step concerns with $\mathbb{E}\,|Y(t)|$ (see Sect. 8). Both steps require sharp estimates and are quite involved (see in particular Sects. 6.2 and 8). Formula (5) implies the bound ($\|\cdot\|_\infty$ stands for the supremum norm)

$$\|\nabla P_t f\|_\infty := \sup_{x\in\mathbb{R}^d} |\nabla P_t f(x)| \le C_T t^{-\frac{1}{\alpha}} \|f\|_\infty, \quad f \in B_b(\mathbb{R}^d), \quad t \in (0,T]. \tag{6}$$

It seems that when $0 < \alpha \le 1$ also estimate (6) is new; it cannot be obtained by a perturbation argument which is available when $\alpha > 1$. In fact we will establish (6) for any process $Z$ with small jumps similar to $\alpha$-stable process. Recall that estimates like (6) for $\alpha > 1$ hold even in some non-degenerate multiplicative cases (see Theorem 1.1 in [15]; in such result the Lipschitz case $\gamma = 1$ requires $\alpha > 1$). We expect that our approach should also work for SDEs with multiplicative cylindrical noise; such an extension is a subject of our ongoing research.

Let us mention that from the analytical point of view we are concerned with the gradient estimates of the solution to the following equation with a non-local operator

$$\begin{aligned}\frac{\partial u}{\partial t}(t,x) &= \langle b(x), \nabla u(t,x)\rangle \\ &\quad + \sum_{j=1}^d \int_{\mathbb{R}} \Big( u(t,x+\xi e_j) - u(t,x) - \chi_{\{|\xi|\le 1\}}\xi \frac{\partial u}{\partial x_j}(t,x)\Big) m_j(\mathrm{d}\xi), \quad t > 0,\end{aligned}$$

$u(0,x) = f(x)$, where $e_j, j = 1,\ldots,d$, is the canonical basis of $\mathbb{R}^d$.

## 2. Main result

Let $Q_t f(x) = \mathbb{E}\, f(Z^x(t))$ be the transition semigroup corresponding to the Lévy proces $Z^x(t) = x + Z(t)$. The proof of the following theorem concerning BEL formulae for $(P_t)$ and $(Q_t)$ is postponed to Sect. 6.

**Theorem 1.** *Let $P = (P_t)$ be given by (1), (2). Assume that:*

(i) *$b \in C^2(\mathbb{R}^d, \mathbb{R}^d)$ has bounded derivatives $\frac{\partial b_i}{\partial \xi_j}$, $\frac{\partial^2 b_i}{\partial \xi_j \partial \xi_k}$, $i,j,k = 1,\ldots,d$.*

(ii) *There is a $\rho > 0$ such that*

$$\liminf_{\varepsilon\downarrow 0} \varepsilon^\rho m_j\{|\xi| \ge \varepsilon\} \in (0,+\infty], \quad j = 1,\ldots,d.$$

(iii) *There exists a $r>0$ such that each $m_j$ restricted to the interval $(-r,r)$ is absolutely continuous with respect to Lebesgue measure. Moreover, the density $\rho_j = \frac{\mathrm{d}m_j}{\mathrm{d}\xi}$ is of class $C^1((-r,r)\setminus\{0\})$ and there exists a $\kappa>1$ such that for all $j$,*

$$\int_{-r}^{r} |\xi|^{\kappa}\rho_j(\xi)\mathrm{d}\xi < +\infty, \tag{7}$$

$$\int_{-r}^{r} |\xi|^{2\kappa}\left(\frac{\rho_j'(\xi)}{\rho_j(\xi)}\right)^2 \rho_j(\xi)\mathrm{d}\xi < +\infty, \tag{8}$$

$$\int_{-r}^{r} |\xi|^{2\kappa-2}\rho_j(\xi)\mathrm{d}\xi < +\infty. \tag{9}$$

*Then there are integrable random fields $Y(t) = (Y_1(t),\ldots,Y_d(t))$, and $Y(t,x) = (Y_1(t,x),\ldots,Y_d(t,x))$, $t>0$, $x\in\mathbb{R}^d$, such that for any $f\in B_b(\mathbb{R}^d)$, $t>0$, $x\in\mathbb{R}^d$,*

$$\nabla Q_t f(x) = \mathbb{E}\, f(Z^x(t))Y(t)$$

*and the Bismut–Elworthy–Li formula (4) for $(P_t)$ holds. Moreover, for any $T>0$ there is an independent of $t\in(0,T]$ and $x$ constant $C$ such that*

$$\mathbb{E}\left(|Y(t)| + |Y(t,x)|\right) \le Ct^{-\frac{\kappa}{\rho}+\frac{1}{2}}, \tag{10}$$

$$\mathbb{E}\,|Y(t) - Y(t,x)| \le Ct^{-\frac{\kappa}{\rho}+\frac{3}{2}}. \tag{11}$$

**Remark 1.** Note that, what is expected, the rate $-\frac{\kappa}{\rho}+\frac{1}{2}$ depends only on the small jumps of $Z$.

**Remark 2.** In fact we have formulae for the fields appearing in Theorem 1. Namely,

$$\begin{aligned} Y_j(t) &= \sum_{k=1}^{d}\left[A_{k,j}(t)D_k^*\mathbf{1}(t) - D_kA_{k,j}(t)\right], \\ Y_j(t,x) &= \sum_{k=1}^{d}\left[A_{k,j}(t,x)D_k^*\mathbf{1}(t) - D_kA_{k,j}(t,x)\right], \end{aligned} \tag{12}$$

where:

- the matrix-valued random fields $A(t) = [A_{k,j}(t)] \in M(d\times d)$ and $A(t,x) = [A_{k,j}(t,x)] \in M(d\times d)$ are given by

$$\begin{aligned} A(t) &= \left[\mathbb{D}Z(t)\right]^{-1}, \\ A(t,x) &= \left[\mathbb{D}X^x(t)\right]^{-1}\nabla X^x(t), \quad \mathbb{P}-\text{a.}s, \end{aligned} \tag{13}$$

  We note that the matrix $A(t)$ is diagonal with entries

$$A_{j,j}(t) = \left(\int_0^t\int_{\mathbb{R}} V_j(s,\xi_j)\Pi_j(\mathrm{d}s,\mathrm{d}\xi_j)\right)^{-1}.$$

- $\mathbb{D}Z(t)$ and $\mathbb{D}X^x(t)$ are the Malliavin derivatives (see Sect. 3 and formulae (27) and (29)) of $Z(t)$ and $X^x(t)$ respectively, with respect to the field $V = (V_1, \ldots, V_d)$,
$$V_j(t,\xi) = \phi_\delta(\xi_j)\psi_\delta(t) = V_j(t,\xi_j). \tag{14}$$
Here $\psi_\delta \in C^\infty(\mathbb{R})$ and $\phi_\delta \in C^\infty(\mathbb{R}\backslash\{0\})$ are non-negative functions such that
$$\psi_\delta(z) = \begin{cases} 0 & \text{if } |z| \geq \delta, \\ 1 & \text{if } |z| \leq \frac{\delta}{2} \end{cases}, \qquad \phi_\delta(z) = |z|^\kappa \psi_\delta(z), \tag{15}$$
with $\kappa$ appearing in assumption $(iii)$ of Theorem 1, and
$$\delta \in (0,r] \quad \text{small enough.}$$
- $\nabla X^x(t)$ is the derivative in probability of $X^x$ with respect to the initial condition $x$,
- $D_k^*\mathbf{1}(t)$ is the adjoint derivative operator calculated on the constant function $\mathbf{1}$, see Sect. 3, Lemma 3.

**Remark 3.** The fields $Y(t)$ and $Y(t,x)$ are not uniquely determined by the BEL formulae. In particular the BEL formula for $(Q_t)$ holds with $Y(t)$ being replaced by $Y(t)+\eta(t)$, where $\eta(t)$ is any zero-mean random variable which is independent of $Z^x(t)$. Note that the conditional expectations $\mathbb{E}\left(Y(t)|Z^x(t)\right)$ and $\mathbb{E}\left(Y(t,x)|X^x(t)\right)$ are uniquely determined. On the other hand, $\mathbb{E}\,|Y(t)|$ and $\mathbb{E}\,|Y(t,x)|$ may depend on the choice of the fields.

Estimate (10) implies new uniform gradient estimates
$$\|\nabla P_t f\|_\infty \leq C_T t^{-\frac{\kappa}{\rho}+\frac{1}{2}} \|f\|_\infty, \qquad t \in (0,T], \quad f \in B_b(\mathbb{R}^d). \tag{16}$$
Although (16) is quite general, it is not sharp in the relevant cylindrical $\alpha$-stable case with $\alpha \in (0,2)$. In such case $\rho = \alpha$ and $\kappa$ is any real number satisfying $\kappa > 1 + \frac{\alpha}{2}$. Therefore we only get that for any $\varepsilon > 0$ and $T < +\infty$ there is a constant $C_{\varepsilon,T}$ such that for any $f \in B_b(\mathbb{R}^d)$,
$$\|\nabla P_t f\|_\infty \leq C_{\varepsilon,T} t^{-\frac{1}{\alpha}-\varepsilon} \|f\|_\infty, \qquad t \in (0,T]. \tag{17}$$
We will improve the previous estimate in Sect. 8 by considering $\varepsilon = 0$. To this purpose we will also use the next remark.

**Remark 4.** Our main theorem provides also estimate (11) for $\mathbb{E}\,|Y(t,x) - Y(t)|$. This can be useful. Indeed if for some specific Lévy processes $Z_j$ we have
$$\mathbb{E}\,|Y(t)| \leq C_T t^{-\eta}, \qquad t \in (0,T] \tag{18}$$
or even if $\mathbb{E}\,|\mathbb{E}\left(Y(t)|X^x(t)\right)| \leq C_T t^{-\eta}$ for some $\eta$ such that
$$\frac{\kappa}{\rho} - \frac{3}{2} \leq \eta \leq \frac{\kappa}{\rho} - \frac{1}{2},$$
where $\kappa$ verifies our assumptions, then we can improve (10) and get, for $t \in (0,T]$,
$$\mathbb{E}\,|Y(t,x)| \leq C'_T t^{-\eta}, \qquad t \in (0,T]. \tag{19}$$

By (19) one deduces

$$\|\nabla P_t f\|_\infty \le C'_T t^{-\eta} \|f\|_\infty , \qquad t \in (0,T].$$

In particular when $Z_j$ are independent real $\alpha$-stable processes, $\alpha \in (0,2)$, we will get in Sect. 8 the crucial estimate

$$\mathbb{E}\,|Y(t)| \le C_T t^{-\frac{1}{\alpha}}, \qquad t \in (0,T]. \tag{20}$$

Combining (11) with (20) we deduce in the cylindrical $\alpha-$stable case

$$\mathbb{E}\,|Y(t,x)| \le C'_T t^{-\frac{1}{\alpha}}, \qquad t \in (0,T], \tag{21}$$

(where $C'_T$ is independent of $x$ and $t$) and the sharp gradient estimate

$$\|\nabla P_t f\|_\infty \le C'_T t^{-\frac{1}{\alpha}} \|f\|_\infty , \qquad t \in (0,T].$$

**Remark 5.** The time dependent case could be also considered. This is the case when the drift $b(x)$ is replaced by $b(t,x)$ (assuming that $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ is Borel and verifies $|b(t,x)| \le C(1+|x|)$, $b(t,\cdot) \in C^2(\mathbb{R}^d,\mathbb{R}^d)$ with all spatial derivates bounded uniformly in $t \in [0,T]$). In such situation one deals with a time dependent Markov semigroup $(P_{st})$. Fixing $s \in [0,T)$ one could obtain a formula for $\nabla(P_{st}f)(x)$ with $s < t \le T$, $f \in \mathcal{B}_b(\mathbb{R}^d)$ which generalizes (4). The strategy is basically the same as in this paper but the computations would be much more involved.

As mentioned in the introduction a difficulty of the Proof of Theorem 1 is also to show that the Malliavin derivative of the solution $\mathbb{D}(X^x(t))$ in the direction to a suitable random field $V$ is invertible and the inverse is integrable with sufficiently large power. The idea (see the proof of our Lemma 5) is to show that $\mathbb{D}(X^x(t)) \approx \mathbb{D}Z(t)$, where $\mathbb{D}Z(t)$, is a diagonal matrix with the terms $\int_0^t \int_{\mathbb{R}} V_j(s,\xi_j)\Pi_j(\mathrm{d}s,\mathrm{d}\xi_j)$ on diagonal. Therefore the integrability of $\left(\mathbb{D}(X^x(t))\right)^{-1}$ follows from the known fact, see Sect. 5 that

$$\mathbb{E}\left[\int_0^t \int_{\mathbb{R}} V_j(s,\xi_j)\Pi_j(\mathrm{d}s,\mathrm{d}\xi_j)\right]^{-q} \le C(q,T)\, t^{-\frac{\kappa q}{\rho}}, \qquad \forall\, q \in (1,+\infty).$$

On the other hand, several technical difficulties arise in proving the sharp bounds for $\mathbb{E}\,|Y(t) - Y(t,x)|$ and $\mathbb{E}\,|Y(t)|$.

Finally, we mention that an attempt to prove (4) has been done in [4] by the martingale approach used in [21] (see, in particular, Lemma A.3 in [4]). However the BEL formula in [4] does not seem to be correct, since there is a gap in the proof, passing from formula (48) to (49) in page 1450 of [4], which consists in an undue application of the chain rule. It seems that the complication here is substantial, and it is difficult to adapt directly the approach used in [21] to the current setting, where because of singularity of the noise it is hard to guarantee invertibility of the Malliavin derivative w.r.t. one vector field. Exactly this crucial point is our reason to use a matrix-valued Malliavin derivative of the solution w.r.t. a vector-valued field $V = (V_1, \dots, V_d)$.

## 3. Malliavin calculus

In this section we adopt in a very direct way the classical concepts and results of Bass and Cranston [3] and Norris [17] to the case of $Z = (Z_1, \dots, Z_d)^*$ being a Lévy process in $\mathbb{R}^d$ with independent coordinates $Z_j$. For more information on Malliavin calculus for jump processes we refer the reader to the book of Ishikawa [12] (see also [2] and the references therein).

We assume that $Z = (Z_1, \dots, Z_d)^*$ is defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. By the Lévy–Itô decomposition

$$Z(t) = \int_0^t \int_{\mathbb{R}^d} \xi \, \overline{\Pi}(\mathrm{d}s, \mathrm{d}\xi),$$

where $\Pi$ is the Poisson random measure on $E := [0, +\infty) \times \mathbb{R}^d$ with intensity measure $\mathrm{d}s\mu(\mathrm{d}\xi)$,

$$\begin{aligned} \overline{\Pi}(\mathrm{d}s, \mathrm{d}\xi) &:= \widehat{\Pi}(\mathrm{d}s, \mathrm{d}\xi)\chi_{\{|\xi| \le 1\}} + \Pi(\mathrm{d}s, \mathrm{d}\xi)\chi_{\{|\xi|>1\}}, \\ \widehat{\Pi}(\mathrm{d}s, \mathrm{d}\xi) &:= \Pi(\mathrm{d}s, \mathrm{d}\xi) - \mathrm{d}s\mu(\mathrm{d}\xi). \end{aligned}$$

Moreover, as the coordinates of $Z$ are independent,

$$\begin{aligned} \mu(\mathrm{d}\xi) &:= \sum_{j=1}^d \mu_j(\mathrm{d}\xi), \\ \mu_j(\mathrm{d}\xi) &:= \delta_0(\mathrm{d}\xi_1) \dots \delta_0(\mathrm{d}\xi_{j-1}) m_j(\mathrm{d}\xi_j) \delta_0(\mathrm{d}\xi_{j+1}) \dots \delta_0(\mathrm{d}\xi_d), \end{aligned} \tag{22}$$

where $\delta_0$ is the Dirac $\delta$-function, and $m_j(\mathrm{d}\xi_j)$ is the Lévy measure of $Z_j$. Note that

$$\Pi(\mathrm{d}s, \mathrm{d}\xi) = \sum_{j=1}^d \Pi_j(\mathrm{d}s, \mathrm{d}\xi),$$

where $\Pi_j$ are independent Poisson random measures each on $[0, +\infty) \times \mathbb{R}^d$ with the intensity measure $\mathrm{d}s\mu_j$ (we use the same symbol as for the one-dimensional $\Pi_j(\mathrm{d}s, \mathrm{d}\xi)$ appearing in (3) when no confusion may arise).

Consider the filtration

$$\mathfrak{F}_t = \sigma\left(\Pi([0, s] \times \Gamma) \colon 0 \le s \le t, \ \Gamma \in \mathcal{B}(\mathbb{R}^d)\right), \qquad t \ge 0.$$

The Poisson random measure $\Pi$ can be treated as a random element in the space $\mathbb{Z}_+(E)$ of integer-valued measures on $(E, \mathcal{B})$ with the $\sigma$-field $\mathcal{G}$ generated by the family of functions

$$\mathbb{Z}_+(E) \ni \nu \mapsto \nu(A) \in \{0, 1, 2, \dots, +\infty\}, \qquad A \in \mathcal{B}.$$

**Definition 1.** Let $p \in (0, +\infty)$. We call a random variable $\Psi \colon \Omega \mapsto \mathbb{R}$ an $L^p$-functional of $\Pi$ if there is a sequence of bounded measurable functions $\varphi_n \colon \mathbb{Z}_+(E) \mapsto \mathbb{R}$ such that

$$\lim_{n \to +\infty} \mathbb{E}\, |\Psi - \varphi_n(\Pi)|^p = 0. \tag{23}$$

A random variable $\Psi\colon \Omega \mapsto \mathbb{R}$ is called an $L^0$-functional of $\Pi$ if, instead of (23), the convergence in probability holds

$$\varphi_n(\Pi) \stackrel{(\mathbb{P})}{\to} \Psi. \tag{24}$$

The space of all $L^p$-functionals of $\Pi$ is denoted by $L^p(\Pi)$. Note that for $p \geq 1$, $L^p(\Pi)$ is a Banach space with the norm $\|\Psi\|_{L^p(\Pi)} = (\mathbb{E}\,|\Psi|^p)^{1/p}$, and for $p \in (0,1)$, $L^p(\Pi)$ is a Polish space with the metric $\rho_{L^p(\Pi)}(\Phi,\Psi) = \mathbb{E}\,|\Phi-\Psi|^p$.

Assume now that $V=(V_1,\ldots,V_d)\colon [0,+\infty)\times\mathbb{R}^d \mapsto \mathbb{R}^d$ is a field given by (14) and (15). The parameter $\delta$ appearing in (15) will be specified later. Define transformations $\mathcal{Q}^\varepsilon_k$, $\varepsilon>0$ and $k=1,\ldots,d$, $\mathcal{Q}^\varepsilon_k\colon \mathbb{Z}_+(E)\mapsto\mathbb{Z}_+(E)$ as follows

$$\mathcal{Q}^\varepsilon_k\left(\sum_j \delta_{\tau^j,\xi^j}\right) = \sum_j \delta_{\tau^j,\xi^j+\varepsilon V_k(\tau^j,\xi^j_k)e_k},$$

where $e_k, k=1,\ldots,d$, is the canonical basis of $\mathbb{R}^d$.

Now let $\Psi\in L^0(\Pi)$. Write

$$\mathcal{Q}^\varepsilon_k\Psi = (\mathbb{P}) - \lim_{n\to+\infty}\varphi_n(\mathcal{Q}^\varepsilon_k(\Pi)),$$

where $\varphi_n\colon \mathbb{Z}_+(E)\mapsto\mathbb{R}$ are such that (24) holds true. It follows from Lemma 2 below that $\mathcal{Q}^\varepsilon_k\Psi$ is well defined, that is the limit exists and does not depend on the particular choice of an approximation sequence $(\varphi_n)$.

**Definition 2.** We call $\Psi\in L^0(\Pi)$, *differentiable* (with respect to the field $V=(V_1,\ldots,V_d)$) if there exist limits in probability

$$D_k\Psi = (\mathbb{P}) - \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left(\mathcal{Q}^\varepsilon_k(\Psi)-\Psi\right), \quad k=1,\ldots,d.$$

Here $D_k\Psi$ is the *Malliavin derivative* of $\Psi$ along the direction $V_ke_k$.

If $\Psi\in L^0(\Pi)$ is differentiable then we set

$$\mathbb{D}\Psi = (D_1\Psi,\ldots,D_d\Psi).$$

The proof of the following chain rule is standard and left to the reader.

**Lemma 1.** *Assume that $\Psi_1,\ldots,\Psi_m$ are differentiable functionals of $\Pi$. Then for any $f\in C^1_b(\mathbb{R}^m)$ the variable $f(\Psi_1,\ldots,\Psi_m)$ is differentiable and*

$$D_kf(\Psi_1,\ldots,\Psi_m) = \sum_{j=1}^m \frac{\partial f}{\partial x_j}(\Psi_1,\ldots,\Psi_m)\,D_k\Psi_j, \qquad k=1,\ldots,d. \tag{25}$$

Let $\rho_k = \frac{\mathrm{d}m_k}{\mathrm{d}x}$ be the density of the Lévy measure $m_k$ restricted to $(-r,r)\backslash\{0\}\subset\mathbb{R}$. We extend artificially $\rho_k$ putting $\rho_k(0)=1$. Given $\varepsilon\in[-1,1]$ sufficiently small and $k=1,\ldots,d$, define

$$\lambda^\varepsilon_k(t,\xi_k) := \begin{cases} \left(1+\varepsilon\frac{\partial V_k}{\partial\xi_k}(t,\xi_k)\right)\frac{\rho_k(\xi_k+\varepsilon V_k(t,\xi_k))}{\rho_k(\xi_k)} & \text{, if } \xi_k\in\left(-\frac{r}{2},\frac{r}{2}\right)\setminus\{0\},\\ 1 & \text{, otherwise,}\end{cases}$$

$$\Lambda^\varepsilon_k(t,\xi_k) := \lambda^\varepsilon_k(t,\xi_k) - 1 - \log\lambda^\varepsilon_k(t,\xi_k),$$

and

$$M_k^\varepsilon(t) := \exp\left\{\int_0^t \int_{\mathbb{R}^d} \log \lambda_k^\varepsilon(s,\xi_k)\widehat{\Pi}_k(\mathrm{d}s,\mathrm{d}\xi) - \int_0^t \int_{\mathbb{R}^d} \Lambda_k^\varepsilon(s,\xi_k)\mu_k(\mathrm{d}\xi)\mathrm{d}s\right\},$$

where $\mu_k$ is defined in (22) and

$$\widehat{\Pi}_k(\mathrm{d}s,\mathrm{d}\xi) := \Pi_k(\mathrm{d}s,\mathrm{d}\xi) - \mathrm{d}s\mu_k(\mathrm{d}\xi).$$

Note that the set

$$\left\{\xi_k \in \left(-\frac{r}{2},\frac{r}{2}\right) : \xi_k + \varepsilon V_k(t,\xi_k) = 0\right\}$$

is of Lebesgue measure zero.

We will need the following result (see e.g. [17] or [13]).

**Lemma 2.** *The process $M_k^\varepsilon$ is a martingale and for all $T \geq 0$, and $m \in \mathbb{R}$, $\mathbb{E}\left[M_k^\varepsilon(T)\right]^m < +\infty$. Let $T \in (0,+\infty)$. Then, under the probability $\mathrm{d}\mathbb{P}^\varepsilon = M_k^\varepsilon(T)\mathrm{d}\mathbb{P}$, $\mathcal{Q}_k^\varepsilon(\Pi)$ restricted to $[0,T]\times\mathbb{R}^d$ is a Poisson random measure with intensity $\mu_k(\mathrm{d}\xi)\mathrm{d}s$.*

The following lemma provides an integration by parts formula for the derivative $D_k$. For the completeness we repeat some elements of a proof from [17].

**Lemma 3.** *For any $1 \leq q \leq 2$ and $t \in (0,+\infty)$, the random variable*

$$D_k^*\mathbf{1}(t) := -\int_0^t \int_{(-r,r)\times\mathbb{R}^{d-1}} \frac{\frac{\partial}{\partial\xi_k}\left(V_k(s,\xi_k)\rho_k(\xi_k)\right)}{\rho_k(\xi_k)}\widehat{\Pi}_k(\mathrm{d}s,\mathrm{d}\xi) \tag{26}$$

*is $q$-integrable. Assume that $p \geq 2$ and that $\Phi \in L^p(\Pi)$ is differentiable and $\mathfrak{F}_t$-measurable. Then $\mathbb{E}D_k\Phi = \mathbb{E}\Phi D_k^*\mathbf{1}(t)$.*

*Proof.* Note that the process $D_k^*\mathbf{1}(t)$ is well defined and $q$-integrable thanks to (8) and (9). The integrability follows from the fact, see e.g. [1], Theorem 4.4.23, or [19], Lemma 8.22, that one has

$$\mathbb{E}\left|D_k^*\mathbf{1}(t)\right|^2 \leq c\,\mathbb{E}\int_0^t \int_{(-r,r)\times\mathbb{R}} \left|\frac{\frac{\partial}{\partial\xi_k}\left(V_k(s,\xi_k)\rho_k(\xi_k)\right)}{\rho_k(\xi_k)}\right|^2 \mathrm{d}s\rho_k(\xi_k)\mathrm{d}\xi_k.$$

By Lemma 2 we have

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\mathbb{E}\left(\mathcal{Q}_k^\varepsilon\Phi\right)M_k^\varepsilon(t) = 0.$$

Thus

$$0 = \mathbb{E}\left[D_k\Phi M_k^0(t) + \Phi R(t)\right] = \mathbb{E}\left[D_k\Phi + \Phi R(t)\right],$$

where

$$R(t) := \frac{\mathrm{d}}{\mathrm{d}\varepsilon}M_k^\varepsilon(t)|_{\varepsilon=0}.$$

Consequently, we need to show that $D_k^*\mathbf{1}(t) = -R(t)$.

Since

$$M_k^\varepsilon(t) = \exp\left\{\int_0^t \int_{\mathbb{R}^d} \log \lambda_k^\varepsilon(s,\xi_k)\widehat{\Pi}_k(\mathrm{d}s,\mathrm{d}\xi) - \int_0^t \int_{(-r,r)\times\mathbb{R}^{d-1}} \Lambda_k^\varepsilon(s,\xi_k)\mu_k(\mathrm{d}\xi)\mathrm{d}s\right\},$$

we have

$$R(t)=\int_0^t\int_{\mathbb{R}^d}\frac{\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\lambda_k^\varepsilon(s,\xi_k)}{\lambda_k^\varepsilon(s,\xi_k)}|_{\varepsilon=0}\widehat{\Pi}_k(\mathrm{d}s,\mathrm{d}\xi)-\int_0^t\int_{(-r,r)\times\mathbb{R}^{d-1}}\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\Lambda_k^\varepsilon(s,\xi_k)|_{\varepsilon=0}\mu_k(\mathrm{d}\xi)\mathrm{d}s.$$

Finally

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\lambda_k^\varepsilon(s,\xi_k)|_{\varepsilon=0}=\frac{\partial V_k}{\partial\xi_k}(s,\xi_k)+\frac{\rho_k'(\xi_k)}{\rho_k(\xi_k)}V_k(s,\xi_k)=\frac{\frac{\mathrm{d}}{\mathrm{d}\xi_k}\left(V_k(s,\xi_k)\rho_k(\xi_k)\right)}{\rho_k(\xi_k)}.$$

□

## 4. Malliavin derivative of $X^x$

Let $X^x(t)=[X_1^x(t),\dots,X_d^x(t)]^*\in\mathbb{R}^d$ be the value of the solution at time $t$. We use the convention that the vectors in $\mathbb{R}^d$ are columns, and the derivatives (gradients) are rows. Using the chain rule (see Lemma 1) it is easy to check that each of its coordinate is a differentiable functional of $\Pi$ and the $d\times d$-matrix valued process $\mathbb{D}X^x(t)$,

$$[\mathbb{D}X^x(t)]_{i,j}=D_jX_i^x(t)$$

satisfies the following random ODE

$$\mathrm{d}\mathbb{D}X^x(t)=\nabla b(X^x(t))\mathbb{D}X^x(t)\mathrm{d}t+\mathrm{d}Z^V(t),\qquad \mathbb{D}X^x(0)=0 \tag{27}$$

(cf. Section 5 in [3]) where $Z^V(t)=\left[Z_{ij}^V(t)\right]$, $t\geq 0$, is a $d\times d$-matrix valued process

$$Z_{j,j}^V(t):=\int_0^t\int_{\mathbb{R}}V_j(s,\xi)\Pi_j(\mathrm{d}s,\mathrm{d}\xi)=D_jZ_j(t),\qquad Z_{j,i}^V(t)=0\quad\text{if } i\neq j. \tag{28}$$

Note that $\int_{\mathbb{R}}|V_j(t,\xi)|m_j(\mathrm{d}\xi)<+\infty$ thanks to (7), and therefore, the process $Z^V$ is well defined and $q$-integrable for any $q\in[1,+\infty)$. The integrability follows from the so-called Kunita inequality (see [1]) and assumption (7). In fact the Kunita inequality ensures that for $q\geq 2$,

$$\mathbb{E}\left|\int_0^t\int_{\mathbb{R}}V_j(s,\xi_i)\Pi_j(\mathrm{d}s,\mathrm{d}\xi_j)\right|^q\leq C_q\left[\left(\int_0^t\int_{\mathbb{R}}V_j^2(s,\xi_j)\mathrm{d}s\rho_j(\xi_j)\mathrm{d}\xi_j\right)^{q/2}+\int_0^t\int_{\mathbb{R}}V_j^q(s,\xi_j)\mathrm{d}s\rho_j(\xi_j)\mathrm{d}\xi_j\right].$$

Clearly we have:

$$\mathbb{D}Z(t)=Z^V(t),\quad t\geq 0. \tag{29}$$

Let $\nabla X^x(t)$ be the derivative in probability of the solution with respect to the initial value

$$[\nabla X^x(t)]_{i,j}=\frac{\partial}{\partial x_j}X_i^x(t).$$

Note that, the process $X^x$ might not be integrable. However, as the noise is additive and $b$ has bounded derivatives, $\nabla X^x(t)$ exists, it is $p$-integrable, for any $p \geq 1$, and

$$\mathrm{d}\nabla X^x(t) = \nabla b(X^x(t))\nabla X^x(t)\mathrm{d}t, \qquad \nabla X^x(0) = I.$$

Since $b$ has bounded derivatives, we have the next result in which $\|\cdot\|$ is the operator norm on the space of real $d \times d$-matrices.

**Lemma 4.** *For all $t \geq 0$ and $x \in \mathbb{R}^d$, $\nabla X^x(t)$ is an invertible matrix. Moreover, there is a constant $C$ such that*

$$\|\nabla X^x(t)\| + \|\left(\nabla X^x(t)\right)^{-1}\| \leq C\mathrm{e}^{Ct}, \qquad \forall\, t \geq 0,\ \forall\, x \in \mathbb{R}^d.$$

*Moreover, there is a constant $C$, possibly depending on $T$, such that*

$$\|\nabla X^x(t) - I\| + \|\left(\nabla X^x(t)\right)^{-1} - I\| \leq Ct, \qquad \forall\, t \in [0,T],\ \forall\, x \in \mathbb{R}^d.$$

As a simple consequence of (27) and Lemma 4 we have

$$\mathbb{D}X^x(t) = \nabla X^x(t) \int_0^t \left(\nabla X^x(s)\right)^{-1} \mathrm{d}Z^V(s). \tag{30}$$

Let

$$M(t,x) := \int_0^t \left(\nabla X^x(s)\right)^{-1} \mathrm{d}Z^V(s). \tag{31}$$

Then $\mathbb{D}X^x(t) = \nabla X^x(t)M(t,x)$ and consequently the matrix valued process $A = [A_{k,j}(t,x)]$ given by (13) satisfies

$$A(t,x) = \left(\mathbb{D}X^x(t)\right)^{-1} \nabla X^x(t) = \left(M(t,x)\right)^{-1}. \tag{32}$$

The proof of the following lemma is moved to the next section (Sect. 5).

**Lemma 5.** *Assume that the parameter $\delta$ in (15) is small enough. Let $p \geq 1$. The Malliavin matrix $\mathbb{D}X^x(t)$ is invertible and $p$-integrable. Moreover, the matrix valued process $A = [A_{k,j}(t,x)]$ given by (13) or (32) is differentiable and $p$-integrable.*

## 5. Proof of Lemma 5

As before $\|\cdot\|$ denotes

the operator norm on the space of real $d \times d$-matrices. Moreover for a random $d \times d$-matrix $B$ we set

$$\|B\|_{L^p} = \left(\mathbb{E}\,\|B\|^p\right)^{1/p}, \quad p \geq 1.$$

**Lemma 6.** *(i) For any $t > 0$, the matrix $Z^V(t)$ is invertible, $\mathbb{P}$-a.s.. Moreover, for any $p \geq 1$, $T > 0$, there is a constant $C = C(p,T)$ such that*

$$\|\left(Z^V(t)\right)^{-1}\|_{L^p} \leq Ct^{-\frac{\kappa}{\rho}}, \qquad t \in (0,T].$$

*(ii) Assume that the parameter $\delta$ in (15) is small enough (possibly depending on the dimension $d$). Then the matrix $M(t,x)$ is invertible, $\mathbb{P}$-a.s.. Moreover, for any $p \geq 1$ and any $T > 0$, there is a constant $C = C(p,T)$ such that*

$$\|A(t,x) - \left(Z^V(t)\right)^{-1}\|_{L^p} \leq Ct^{-\frac{\kappa}{\rho}+1}, \qquad t \in (0,T] \tag{33}$$

*where $A(t,x) = (M(t,x))^{-1}$.*

*Proof.* The first part of the lemma follows from Corollary 1 from Sect. 7 below. To show the second part note that

$$M(t,x) = Z^V(t) + \int_0^t R(s,x)\mathrm{d}Z^V(s),$$

where $R(t,x) := (\nabla X^x(t))^{-1} - I$ is a random variable taking values in the space of $d \times d$ matrices. Note that

$$\begin{aligned}\left(\int_0^t R(s,x)\mathrm{d}Z^V(s)\right)_{i,j} &= \int_0^t \int_{\mathbb{R}} R_{ij}(s,x)V_j(s,\xi_j)\Pi_j(\mathrm{d}s,\mathrm{d}\xi_j)\\ &= \left(\sum_{0<s\leq t} R(s,x)\tilde{V}(s,\triangle Z(s))\right)_{i,j},\end{aligned}$$

with $\triangle Z(s) = Z(s) - Z(s-)$, where $\tilde{V}(s,z)$ is a diagonal matrix, $s \geq 0$, $z \in \mathbb{R}^d$, such that

$$(\tilde{V}(s,z))_{i,i} = V_i(s,z_i), \quad i = 1,\dots,d.$$

Moreover, $\mathbb{P}$-a.s., $Z^V(t) = \sum_{0<s\leq t} \tilde{V}(s,\triangle Z(s))$ is convergent by (7) and it is also invertible. We write

$$M(t,x) = \left(I + \int_0^t R(s,x)\mathrm{d}Z^V(s)\,(Z^V(t))^{-1}\right) Z^V(t). \tag{34}$$

We would like to obtain, for $\delta > 0$ small enough, $t > 0$,

$$A(t,x) = (Z^V(t))^{-1}\left(I + \int_0^t R(s,x)\mathrm{d}Z^V(s)\,(Z^V(t))^{-1}\right)^{-1}. \tag{35}$$

To this purpose we consider

$$Q(t,x) = \int_0^t R(s,x)\mathrm{d}Z^V(s)\,(Z^V(t))^{-1}.$$

Recall that $(e_j)$ is the canonical basis of $\mathbb{R}^d$. We get for $j = 1,\dots,d$, $\mathbb{P}$-a.s.,

$$\begin{aligned}Q(t,x)e_j &= \sum_{0<s\leq t} R(s,x)\tilde{V}(s,\triangle Z(s))e_j\left(\int_0^t\int_{\mathbb{R}} V_j(y,\xi_j)\Pi_j(\mathrm{d}y,\mathrm{d}\xi_j)\right)^{-1}\\ &= \sum_{0<s\leq t} R(s,x)V_j(s,\triangle Z_j(s))e_j\left(\int_0^t\int_{\mathbb{R}} V_j(y,\xi_j)\Pi_j(\mathrm{d}y,\mathrm{d}\xi_j)\right)^{-1}\end{aligned}$$

and so

$$
\begin{aligned}
|Q(t,x)e_j| &\le \sum_{0<s\le t} \|R(s,x)\|\, V_j(s,\triangle Z_j(s)) \left(\int_0^t \int_{\mathbb{R}} V_j(y,\xi_j)\Pi_j(\mathrm{d}y,\mathrm{d}\xi_j)\right)^{-1} \\
&\le \min\left(Ct,1/2\right) \int_0^t \int_{\mathbb{R}} V_j(y,\xi_j)\Pi_j(\mathrm{d}y,\mathrm{d}\xi_j) \left(\int_0^t \int_{\mathbb{R}} V_j(s,\xi_j)\Pi_j(\mathrm{d}s,\mathrm{d}\xi_j)\right)^{-1} \\
&= \min\left(Ct,1/2\right), \qquad (36)
\end{aligned}
$$

where $C$ is independent of $x\in\mathbb{R}^d$, $t\ge 0$ and $\omega$, $\mathbb{P}$-a.s. Above we used the second estimate of Lemma 4; $\|R(s,x)\|\le Cs$. We will need also that $|Q(t,x)e_j|\le 1/2$. To this end we have to consider $\delta$ sufficiently small. In fact we require $\delta C\le 1/2$.

Therefore, as $Z^V(t)$ is invertible, the matrix $M(t,x)$ is invertible and $A(t,x)=(M(t,x))^{-1}$ satisfies (35). Moreover

$$
A(t,x) = \left(Z^V(t)\right)^{-1} + \left(Z^V(t)\right)^{-1} \sum_{n=1}^{+\infty} (-1)^n (Q(t,x))^n. \qquad (37)
$$

Consequently, we have

$$
\left\| A(t,x) - \left(Z^V(t)\right)^{-1} \right\|_{L^p} \le C_1 t \,\|\left(Z^V(t)\right)^{-1}\|_{L^p}
$$

and (33) follows. The proof is complete. □

**Remark 6.** We note that in the previous proof it is important to have a term like $\int_0^t R(s,x)\mathrm{d}Z^V(s)\,(Z^V(t))^{-1}$ (cf. (34)). Such term can be estimated in a sharp way by $\min(Ct,1/2)$. On the other hand, a term like $(Z^V(t))^{-1}\int_0^t R(s,x)$ $\mathrm{d}Z^V(s)$ would be difficult to estimate in a sharp way (we can estimate its $L^2$-norm by $Ct^{-\frac{\kappa}{\rho}+\frac{3}{2}}$). On this respect see also the computations in Sect. 6.2.

### 5.1. Proof of Lemma 5

Since $b$ has bounded derivatives of the first and second order, $\nabla X^x(t)$ and $(\nabla X^x(t))^{-1}$ are differentiable and $p$-integrable. Next, thanks to (9), the matrix valued process $Z^V$ given by (28) is also differentiable, $p$-integrable, and

$$
\begin{aligned}
D_k Z^V_{k,k}(t) &= \frac{\mathrm{d}}{\mathrm{d}\varepsilon} \int_0^t \int_{\mathbb{R}} V_k(s,\xi_k+\varepsilon V_k(s,\xi_k))\Pi_k(\mathrm{d}s,\mathrm{d}\xi_k)|_{\varepsilon=0} \\
&= \int_0^t \int_{\mathbb{R}} \psi_\delta^2(s)\phi_\delta(\xi_k)\phi_\delta'(\xi_k)\Pi_k(\mathrm{d}s,\mathrm{d}\xi_k).
\end{aligned} \qquad (38)
$$

Therefore, as

$$
\mathrm{d}\mathbb{D}X^x(t) = \nabla b(X^x(t))\mathbb{D}X^x(t)\mathrm{d}t + \mathrm{d}Z^V(t), \qquad \mathbb{D}X^x(0)=0,
$$

$b$ has bounded derivatives of the first and second order, and $\mathrm{d}Z^V(t)$ is $p$-integrable and differentiable, we infer that $\mathbb{D}X^x(t)$ is $p$-integrable and differentiable. Clearly $\nabla X^x(t)$ is invertible. By Lemma 6, the matrix $M(t,x)$ given by (31) is invertible, $p$-integrable and differentiable. Since, (cf. (30) and (31)),

$$
M(t,x) = (\nabla X^x(t))^{-1}\,\mathbb{D}X^x(t)
$$

and, by Lemma 6, $A(t,x):=(M(t,x))^{-1}$ is $p$-integrable, we infer that $\mathbb{D}X^x(t)$ is invertible, and $(\mathbb{D}X^x(t))^{-1}$ is $p$-integrable.

We can show the differentiability of $(\mathbb{D}X^x(t))^{-1}$ or equivalently of $A(t,x)$ in a standard way based on the observation that

$$D_k\left(\mathbb{D}X^x(t)\right)^{-1} = -\left(\mathbb{D}X^x(t)\right)^{-1}\left(D_k\mathbb{D}X^x(t)\right)\left(\mathbb{D}X^x(t)\right)^{-1}. \qquad \square$$

## 6. Proof of Theorem 1

By Lemma 5 the random field $Y(t,x)$ given by (12) is well defined and integrable. By an approximation argument, see e.g. [21], Corollary 3.1 and its proof given in Section 4.3, or [14], see also [18], Lemma 2.2 for gradient estimates, it is enough to show that for any $f \in C_b^1(\mathbb{R}^d)$ we have (4). To this end note that

$$\nabla P_t f(x) = \nabla \mathbb{E}\, f(X^x(t)) = \mathbb{E}\, \nabla f(X^x(t))\nabla X^x(t).$$

Since, by Lemma 1,

$$\mathbb{D}f(X^x(t)) = \nabla f(X^x(t))\mathbb{D}X^x(t),$$

and, by Lemma 5 the matrix $\mathbb{D}X^x(t)$ is invertible, we have

$$\begin{aligned}\nabla P_t f(x) &= \mathbb{E}\left(\mathbb{D}f(X^x(t))\right)\left[\mathbb{D}X^x(t)\right]^{-1}\nabla X^x(t)\\ &= \mathbb{E}\left(\mathbb{D}f(X^x(t))\right)A(t,x) = \sum_{j=1}^d\sum_{k=1}^d \mathbb{E}\, D_k f(X^x(t))A_{k,j}(t,x)e_j^*,\end{aligned}$$

where $A(t,x)$ is given by (13) or equivalently by (32), and, as gradients are row vectors, $e_j^*$ is the transpose of $e_j$. By the chain rule we have

$$\sum_{k=1}^d D_k f(X^x(t))A_{k,j}(t,x) = \sum_{k=1}^d \left\{D_k\left[f(X^x(t))A_{k,j}(t,x)\right] - f(X^x(t))D_kA_{k,j}(t,x)\right\}.$$

Hence, by Lemma 3, we have (4) with $Y$ given by (12). The same arguments can be applied to show the BEL formula for the Lévy semigroup.

The proof of (10) and (11) is more difficult, and it is divided into the following two parts.

### 6.1. Lévy case

Assume that $b \equiv 0$, that is $X^x(t) = Z^x(t)$. Let us fix a time horizon $T < +\infty$. We are proving estimate (10) for the process $Y(t)$ corresponding to the pure Lévy case.

We have, for $j = 1, \ldots, d$,

$$Y_j(t) = \sum_{k=1}^d \left[A_{k,j}(t)D_k^*\mathbf{1}(t) - D_kA_{k,j}(t)\right],$$

where $A(t) = [\mathbb{D}Z^x(t)]^{-1} = \left[Z^V(t)\right]^{-1}$ and $Z^V(t)$ is a diagonal matrix defined in (28). Therefore

$$Y_j(t) = \frac{D_j^*\mathbf{1}(t)}{Z_{j,j}^V(t)} - D_j\frac{1}{Z_{j,j}^V(t)} = \frac{D_j^*\mathbf{1}(t)}{Z_{j,j}^V(t)} + \frac{D_jZ_{j,j}^V(t)}{\left(Z_{j,j}^V(t)\right)^2},$$

where $D_j^*\mathbf{1}(t)$ and $D_j Z^V_{j,j}$ are given by (26) and (38), respectively. We have

$$\mathbb{E}\left|D_j^*\mathbf{1}(t)\left(Z^V_{j,j}(t)\right)^{-1}\right| \le \left(\mathbb{E}\left|D_j^*\mathbf{1}(t)\right|^2\right)^{\frac{1}{2}}\left(\mathbb{E}\left|Z^V_{j,j}(t)\right|^{-2}\right)^{\frac{1}{2}}.$$

By Lemma 6, there is a constant $C_1$ such that $\mathbb{E}\left|Z^V_{jj}(t)\right|^{-2} \le C_1 t^{-\frac{2\kappa}{\rho}}$. Next there are constants $C_2$ and $C_3$ such that

$$\mathbb{E}\left|D_j^*\mathbf{1}(t)\right|^2 \le C_2\int_0^t\int_{-\delta}^{\delta}\left|\frac{\frac{\partial}{\partial\xi_j}\left(V_j(s,\xi_j)\rho_j(\xi_j)\right)}{\rho_j(\xi_j)}\right|^2\rho_j(\xi_j)\mathrm{d}\xi_j\mathrm{d}s \le C_3 t, \quad (39)$$

where the last estimate follows from (8) and (9). Therefore there is a constant $C_4$ such that

$$\mathbb{E}\left|D_j^*\mathbf{1}(t)\left(Z^V_{j,j}(t)\right)^{-1}\right| \le C_4 t^{-\frac{\kappa}{\rho}+\frac{1}{2}}, \qquad t\in(0,T]. \quad (40)$$

Let us observe now that

$$\begin{aligned}
&\left|D_j Z^V_{j,j}(t)\right| = \left|\int_0^t\int_{\mathbb{R}}\psi_\delta^2(s)\phi_\delta(\xi_j)\phi_\delta'(\xi_j)\Pi_j(\mathrm{d}s,\mathrm{d}\xi_j)\right|\\
&\le\left(\int_0^t\int_{\mathbb{R}}\psi_\delta^2(s)\phi_\delta^2(\xi_j)\Pi_j(\mathrm{d}s,\mathrm{d}\xi_j)\right)^{1/2}\left(\int_0^t\int_{\mathbb{R}}\psi_\delta^2(s)\left(\phi_\delta'(\xi_j)\right)^2\Pi_j(\mathrm{d}s,\mathrm{d}\xi_j)\right)^{1/2}\\
&\le\int_0^t\int_{\mathbb{R}}\psi_\delta(s)\phi_\delta(\xi_j)\Pi_j(\mathrm{d}s,\mathrm{d}\xi_j)\left(\int_0^t\int_{\mathbb{R}}\psi_\delta^2(s)\left(\phi_\delta'(\xi_j)\right)^2\Pi_j(\mathrm{d}s,\mathrm{d}\xi_j)\right)^{1/2};
\end{aligned}$$

here in the last inequality we have used an elementary inequality

$$\sum_k x_k^2 \le \left(\sum_k x_k\right)^2,$$

valid for any non-negative real numbers $\{x_k\}$. Thus

$$\left|D_j Z^V_{j,j}(t)\right| \le Z^V_{j,j}(t)\left(\int_0^t\int_{\mathbb{R}}\psi_\delta^2(s)\left(\phi_\delta'(\xi_j)\right)^2\Pi_j(\mathrm{d}s,\mathrm{d}\xi_j)\right)^{1/2}. \quad (41)$$

Therefore, by Lemma 6,

$$\begin{aligned}
\mathbb{E}\left|\frac{D_j Z^V_{j,j}(t)}{\left(Z^V_{j,j}(t)\right)^2}\right| &\le \mathbb{E}\,\frac{\left(\int_0^t\int_{\mathbb{R}}\psi_\delta^2(s)\left(\phi_\delta'(\xi_j)\right)^2\Pi_j(\mathrm{d}s,\mathrm{d}\xi_j)\right)^{1/2}}{\int_0^t\int_{\mathbb{R}}\psi_\delta(s)\phi_\delta(\xi_j)\Pi_j(\mathrm{d}s,\mathrm{d}\xi_j)}\\
&\le\left(\mathbb{E}\int_0^t\int_{\mathbb{R}}\psi_\delta^2(s)\left(\phi_\delta'(\xi_j)\right)^2\Pi_j(\mathrm{d}s,\mathrm{d}\xi_j)\right)^{1/2}\\
&\quad\times\left(\mathbb{E}\left(\int_0^t\int_{\mathbb{R}}\psi_\delta(s)\phi_\delta(\xi_j)\Pi_j(\mathrm{d}s,\mathrm{d}\xi_j)\right)^{-2}\right)^{1/2} \le C_5 t^{-\frac{\kappa}{\rho}+\frac{1}{2}}.
\end{aligned} \quad (42)$$

Note that $\int_{\mathbb{R}}\left(\phi_\delta'(\xi_j)\right)^2 m_j(\mathrm{d}\xi_j) < +\infty$ thanks to (9). Summing up, we can find a constant $C$ such that

$$\mathbb{E}\left|Y(t)\right| \le C t^{-\frac{\kappa}{\rho}+\frac{1}{2}}, \quad (43)$$

which is the desired estimate. □

### 6.2. General case

Recall that $M$ and $A = M^{-1}$ are given by (31) and (32), respectively. Let $T > 0$. We prove first that (for $\delta > 0$ small enough) there is a constant $c$ such that for $t \in (0, T]$,

$$\mathbb{E}\left|D_k^*\mathbf{1}(t)A_{k,j}(t,x)\right| \le ct^{-\frac{\kappa}{\rho}+\frac{1}{2}}, \tag{44}$$

$$\mathbb{E}\left|D_k^*\mathbf{1}(t)A_{k,j}(t,x) - D_k^*\mathbf{1}(t)\left(Z^V(t)\right)^{-1}_{k,j}\right| \le ct^{-\frac{\kappa}{\rho}+\frac{3}{2}}. \tag{45}$$

By Lemma 6 there is a constant $C > 0$ such that

$$\|A(t,x) - \left(Z^V(t)\right)^{-1}\|_{L^q} \le Ct^{-\frac{\kappa}{\rho}+1}, \quad t \in (0,T], \quad q \ge 1.$$

Therefore, (45) follows from (39) by using the Cauchy–Schwarz inequality. Clearly (44) follows from (40) and (45).

It is much harder to evaluate $L^1$-norm of the term

$$I(t,x) := -\sum_{j=1}^d\sum_{k=1}^d D_kA_{k,j}(t,x)e_j = \sum_{j=1}^d\sum_{k=1}^d \left[A(t,x)(D_kM(t,x))A(t,x)\right]_{k,j} e_j. \tag{46}$$

Recall that $R(s,x) := (\nabla X^x(s))^{-1} - I$. Moreover,

$$M(t,x) = Z^V(t) + \int_0^t R(s,x)\mathrm{d}Z^V(s)$$

is differentiable, $p$-integrable, and we have (see also (38)):

$$D_kM(t,x) = D_kZ^V(t) + \int_0^t R(s,x)\mathrm{d}D_kZ^V(s) + \int_0^t D_kR(s,x)\mathrm{d}Z^V(s). \tag{47}$$

We have $\|R(s,x)\| \le C_1 s$, $s \in [0,T]$, and that there are non-negative random variables $\eta(s)$, integrable with an arbitrary power, such that, $\mathbb{P}$-a.s., $0 \le \eta(s) \le \eta(t)$, $0 \le s \le t \le T$,

$$\|D_kR(s,x)\| \le \eta(s), \qquad \|\eta(s)\|_{L^2} \le C_2\, s^{\frac{3}{2}}, \qquad s \in [0,T], \tag{48}$$

where $C_2$ is independent of $s$. Indeed, using that $\mathrm{d}\nabla X^x(t) = \nabla b(X^x(t))\nabla X^x(t)$ $\mathrm{d}t$, $\nabla X^x(0) = I$,

$$\mathrm{d}R(t,x) = -\left[R(t,x)\nabla b(X^x(t)) + \nabla b(X^x(t))\right]\mathrm{d}t, \qquad R(0,x) = 0.$$

Since $\nabla b$ is bounded we have $\|R(s,x)\| \le C_1 s$. After differentiation we obtain

$$\mathrm{d}D_kR(t,x) = -\left[D_kR(t,x)\nabla b(X^x(t)) + (R(t,x)+1)\sum_{i=1}^d \frac{\partial}{\partial x_i}\nabla b(X^x(t))D_kX_i^x(t)\right]\mathrm{d}t,$$
$$D_kR(0,x) = 0.$$

By (30), there is a constant $C_3$ such that for all $t \in [0,T]$, $\|D_k X^x(t)\| \leq C_3\|Z^V(t)\|$. Therefore there is a constant $C_4$ such that

$$\|D_k R(t,x)\| \leq C_4 \int_0^t \left[\|D_k R(s,x)\| + \|Z^V(s)\|\right] \mathrm{d}s,$$

and consequently

$$\|D_k R(t,x)\| \leq \eta(t) := C_5 \int_0^t \|Z^V(s)\| \mathrm{d}s, \quad t \in [0,T].$$

We will show that $I(t,x)$ is a proper perturbation of the already estimated

$$I_0(t) := \sum_{j=1}^d \frac{D_j Z^V_{j,j}(t)}{\left(Z^V_{j,j}(t)\right)^2} e_j = \sum_{j=1}^d \sum_{k=1}^d \left[\left(Z^V(t)\right)^{-1} \left(D_k Z^V(t)\right) \left(Z^V(t)\right)^{-1}\right]_{k,j} e_j. \tag{49}$$

The proof will be completed as soon as we can show there is a constant $C_6$ such that

$$\mathbb{E}\,|I(t,x) - I_0(t)| \leq C_6 t^{-\frac{\kappa}{\rho}+\frac{3}{2}}, \qquad t \in (0,T]. \tag{50}$$

This will imply that

$$\mathbb{E}\,|I(t,x)| \leq \mathbb{E}\,|I(t,x) - I_0(t)| + \mathbb{E}\,|I_0(t)| \leq C_7 t^{-\frac{\kappa}{\rho}+\frac{1}{2}}. \tag{51}$$

Collecting (44) and (51) will give the estimate for $\mathbb{E}\,|Y(t,x)|$.

Let us prove (50). Recalling that $A(t,x) = (M(t,x))^{-1}$ we have to estimate

$$\|A(t,x)(D_k M(t,x))A(t,x) - (Z^V(t))^{-1}(D_k Z^V(t))\left(Z^V(t)\right)^{-1}\| \leq J_1 + J_2 + J_3,$$

where

$$\begin{aligned} J_1 &= \|A(t,x)(D_k M(t,x))[A(t,x) - \left(Z^V(t)\right)^{-1}]\|, \\ J_2 &= \|A(t,x)(D_k M(t,x) - D_k Z^V(t))\,(Z^V(t))^{-1}\|, \\ J_3 &= \|[A(t,x) - (Z^V(t))^{-1}]\,D_k Z^V(t)\,(Z^V(t))^{-1}\|. \end{aligned}$$

We have

$$\mathbb{E}\,[J_3] \leq \|[A(t,x) - (Z^V(t))^{-1}]\|_{L^2}\,\|D_k Z^V(t)\,(Z^V(t))^{-1}\|_{L^2}.$$

Using (33) we infer

$$\mathbb{E}\,[J_3] \leq C_8 t^{-\frac{\kappa}{\rho}+1}\|D_k Z^V(t)\,(Z^V(t))^{-1}\|_{L^2}.$$

Since

$$\|D_k Z^V(t)\,(Z^V(t))^{-1}\| = \left|\frac{D_k Z^V_{k,k}(t)}{Z^V_{k,k}(t)}\right|,$$

we can use (41) and get

$$\mathbb{E}\left\|D_k Z^V(t)\,(Z^V(t))^{-1}\right\|^2 \leq \mathbb{E}\left(\int_0^t \int_{\mathbb{R}} \psi_\delta^2(s)\left(\phi'_\delta(\xi_k)\right)^2 \Pi_k(\mathrm{d}s,\mathrm{d}\xi_k)\right) \leq C_9 t,$$

see (9). Hence we have

$$\mathbb{E}\left[J_3\right] \le C_{10} t^{-\frac{\kappa}{\rho}+\frac{3}{2}}.$$

We evaluate now $J_2$. By Lemma 6 we have

$$\begin{aligned}\mathbb{E}\left[J_2\right] &\le \|A(t,x)\|_{L^2} \left\|\left(D_k M(t,x) - D_k Z^V(t)\right)\left(Z^V(t)\right)^{-1}\right\|_{L^2}\\ &\le C t^{-\frac{\kappa}{\rho}} \left\|\left(D_k M(t,x) - D_k Z^V(t)\right)\left(Z^V(t)\right)^{-1}\right\|_{L^2}.\end{aligned}$$

Next

$$\begin{aligned}&\left(D_k M(t,x) - D_k Z^V(t)\right)\left(Z^V(t)\right)^{-1}\\ &= \left(\int_0^t R(s,x)\mathrm{d}D_k Z^V(s) + \int_0^t D_k R(s,x)\mathrm{d}Z^V(s)\right)\left(Z^V(t)\right)^{-1}.\end{aligned} \tag{52}$$

We will argue as in the proof of Lemma 6. Note that

$$\left(\int_0^t R(s,x)\mathrm{d}D_k Z^V(s)\right)_{ij} = 0 \quad \text{if } j \ne k.$$

Recall that, for $\delta$ small enough,

$$\begin{aligned}\left(\int_0^t R(s,x)\mathrm{d}D_k Z^V(s)\right)_{ik} &= \int_0^t \psi_\delta^2(s) R_{ik}(s,x) \int_{\mathbb{R}} \phi_\delta(\xi_k)\phi_\delta'(\xi_k)\Pi_k(\mathrm{d}s,\mathrm{d}\xi_k)\\ &= \left(\sum_{0<s\le t} R(s,x)\tilde{U}(s,\triangle Z(s))\right)_{ik},\end{aligned}$$

where $\tilde{U}(s,z)$ is a diagonal matrix, $s \ge 0$, $z \in \mathbb{R}^d$, such that

$$(\tilde{U}(s,z))_{ii} = U_i(s,z) = \psi_\delta^2(s)\phi_\delta(z_i)\phi_\delta'(z_i), \quad i = 1,\ldots,d.$$

Hence

$$\begin{aligned}&\Big|\int_0^t R(s,x)\mathrm{d}D_k Z^V(s)\left(Z^V(t)\right)^{-1} e_k\Big|\\ &= \Big|\sum_{0<s\le t} R(s,x)\tilde{U}(s,\triangle Z(s)) e_k \Big(\int_0^t \int_{\mathbb{R}} V_k(y,\xi_k)\Pi_k(\mathrm{d}y,\mathrm{d}\xi_k)\Big)^{-1}\Big|\\ &\le \sum_{0<s\le t} \|R(s,x)\| U_k(s,\triangle Z_k(s)) \Big(\int_0^t \int_{\mathbb{R}} V_k(y,\xi_k)\Pi_k(\mathrm{d}y,\mathrm{d}\xi_k)\Big)^{-1}\\ &\le C_1 t \sum_{0<s\le t} U_k(s,\triangle Z_k(s)) \Big(\int_0^t \int_{\mathbb{R}} V_k(y,\xi_k)\Pi_k(\mathrm{d}y,\mathrm{d}\xi_k)\Big)^{-1} = C_1 t\, \frac{D_k Z^V_{k,k}(t)}{Z^V_{k,k}(t)},\end{aligned}$$

see (38). We deduce that

$$\Big\|\int_0^t R(s,x)\mathrm{d}D_k Z^V(s)\left(Z^V(t)\right)^{-1}\Big\|_{L^2} \le C_{11} t^{\frac{3}{2}}.$$

Therefore, in order to estimate $J_2$, it remains to consider

$$\int_0^t D_k R(s,x)\mathrm{d}Z^V(s)\left(Z^V(t)\right)^{-1} = \sum_{0<s\le t} D_k R(s,x)\tilde{V}(s,\triangle Z(s))(Z^V(t))^{-1},$$

where $\tilde{V}(s,z)$ is a diagonal matrix, $s \geq 0$, $z \in \mathbb{R}^d$, such that $(\tilde{V}(s,z))_{ii} = V_i(s,z_i)$. Using the bound (48) we obtain, for $j = 1,\ldots,d$, $\mathbb{P}$-a.s.,

$$
\begin{aligned}
&\Big| \int_0^t D_k R(s,x) \mathrm{d}Z^V(s) \; \left(Z^V(t)\right)^{-1} e_j \Big| \\
&\quad \leq \sum_{0<s\leq t} \|D_k R(s,x)\| \, V_j(s, \triangle Z_j(s)) \left( \int_0^t \int_{\mathbb{R}} V_j(r,\xi_j) \Pi_j(\mathrm{d}r, \mathrm{d}\xi_j) \right)^{-1} \\
&\quad \leq \eta(t).
\end{aligned}
$$

It follows that

$$
\Big\| \int_0^t D_k R(s,x) \mathrm{d}Z^V(s) \; \left(Z^V(t)\right)^{-1} \Big\|_{L^2} \leq C_{12} \, t^{\frac{3}{2}}.
$$

Summing up we have

$$
\mathbb{E}\,[J_2] \leq C_{13} t^{-\frac{\kappa}{\rho} + \frac{3}{2}}.
$$

To treat $J_1$ we note that by Lemma 6 we have

$$
\begin{aligned}
\mathbb{E}\,[J_1] &\leq \|A(t,x)\|_{L^2} \|(D_k M(t,x))[A(t,x) - \left(Z^V(t)\right)^{-1}]\|_{L^2} \\
&\leq C t^{-\frac{\kappa}{\rho}} \, \|(D_k M(t,x))[A(t,x) - \left(Z^V(t)\right)^{-1}]\|_{L^2}.
\end{aligned}
\tag{53}
$$

We write

$$
\begin{aligned}
&\|(D_k M(t,x))[A(t,x) - \left(Z^V(t)\right)^{-1}]\| \\
&\quad = \|(D_k M(t,x))(M(t,x))^{-1} M(t,x)[A(t,x) - \left(Z^V(t)\right)^{-1}]\| \\
&\quad \leq \|(D_k M(t,x))(M(t,x))^{-1}\| \, \|I - M(t,x) \left(Z^V(t)\right)^{-1} \|.
\end{aligned}
$$

The more difficult term is

$$
\begin{aligned}
\|(D_k M(t,x))(M(t,x))^{-1}\| &= \|(D_k M(t,x)) \left(Z^V(t)\right)^{-1} \; Z^V(t)(M(t,x))^{-1}\| \\
&\leq \|(D_k M(t,x)) \left(Z^V(t)\right)^{-1} \| \, \|Z^V(t)(M(t,x))^{-1}\|.
\end{aligned}
$$

By (37) we have

$$
\begin{aligned}
Z^V(t) \, (M(t,x))^{-1} &= Z^V(t) A(t,x) \\
&= Z^V(t) \left( \left(Z^V(t)\right)^{-1} + \left(Z^V(t)\right)^{-1} \sum_{n=1}^{+\infty} (-1)^n (Q(t,x))^n \right) \\
&= \sum_{n=0}^{+\infty} (-1)^n (Q(t,x))^n.
\end{aligned}
$$

Hence, by (36),

$$
\| \left(Z^V(t)\right) (M(t,x))^{-1} \| \leq \tilde{C}_1,
$$

where $\tilde{C}_1$ is independent of $x$, $t \in (0,T]$ and $\omega$, $\mathbb{P}$-a.s.. The term

$$
\|(D_k M(t,x)) \left(Z^V(t)\right)^{-1} \|
$$

can be treated as the first term in (52). Therefore we have

$$\|(D_k M(t,x))\left(Z^V(t)\right)^{-1}\|_{L^2} \le \tilde{C}_2 t^{\frac{1}{2}}.$$

Summing up we have

$$\|(D_k M(t,x))(M(t,x))^{-1}\|_{L^2} \le \tilde{C}_3 t^{\frac{1}{2}}, \quad t \in (0,T].$$

Since

$$\|I - M(t,x)\left(Z^V(t)\right)^{-1}\| = \left\| \int_0^t R(s,x)\mathrm{d}Z^V(s)\ \left(Z^V(t)\right)^{-1} \right\| \le \tilde{C}_4 t,$$

where $c_3$ is independent of $x$ and $\omega$, $\mathbb{P}$-a.s., we have

$$\mathbb{E}\left[J_1\right] \le \tilde{C}_5 t^{-\frac{\kappa}{\rho}+\frac{3}{2}},$$

and the proof is complete. □

## 7. An integrability result

Assume that $\mathcal{M}$ is a Poisson random measure on $[0,+\infty) \times \mathbb{R}$ with intensity measure $\mathrm{d}t m(\mathrm{d}\xi)$. Given a measurable $h\colon \mathbb{R} \mapsto [0,+\infty)$ let

$$J_h(t) := \int_0^t \int_{\mathbb{R}} h(\xi)\mathcal{M}(\mathrm{d}s,\mathrm{d}\xi).$$

Then for any $\beta > 0$,

$$\mathbb{E}\,\mathrm{e}^{-\beta J_h(t)} = \exp\left\{-t\int_{\mathbb{R}}\left(1-\mathrm{e}^{-\beta h(\xi)}\right) m(\mathrm{d}\xi)\right\}.$$

Using the identity

$$y^{-q} = \frac{1}{\Gamma(q)}\int_0^{+\infty} \beta^{q-1}\mathrm{e}^{-\beta y}\mathrm{d}\beta, \qquad y > 0,$$

we obtain

$$\begin{aligned}\mathbb{E}\,J_h(t)^{-q} &= \frac{1}{\Gamma(q)}\int_0^{+\infty} \beta^{q-1}\mathbb{E}\,\mathrm{e}^{-\beta J_h(t)}\mathrm{d}\beta\\ &= \frac{1}{\Gamma(q)}\int_0^{+\infty} \beta^{q-1}\exp\left\{-t\int_{\mathbb{R}}\left(1-\mathrm{e}^{-\beta h(\xi)}\right) m(\mathrm{d}\xi)\right\}\mathrm{d}\beta.\end{aligned}$$

Using this method one can obtain (see Norris [17]) the following result.

**Lemma 7.** *If for a certain $\rho > 0$,*

$$\liminf_{\varepsilon\downarrow 0} \varepsilon^{\rho} m\{h \ge \varepsilon\} > 0,$$

*then*

$$\mathbb{E}\,J_h(t)^{-q} \le C t^{-\frac{q}{\rho}}, \qquad q \ge 1,\ t \in (0,1].$$

Let $\phi_\delta \in C^\infty(\mathbb{R}\setminus\{0\})$ be given by (15). Assume that $m(\mathrm{d}\xi)$ satisfies hypothesis $(ii)$ of Theorem 1 and $h=\phi_\delta$. Then

$$\begin{aligned}\liminf_{\varepsilon\downarrow 0}\varepsilon^{\frac{\rho}{\kappa}} m\{\phi_\delta\ge\varepsilon\} &\ge \liminf_{\varepsilon\downarrow 0}\varepsilon^{\frac{\rho}{\kappa}} m\left\{\xi\in\left[-\frac{\delta}{2},\frac{\delta}{2}\right] : |\xi|^\kappa\ge\varepsilon\right\}\\ &\ge \liminf_{\varepsilon\downarrow 0}\varepsilon^{\frac{\rho}{\kappa}} m\left\{\xi\colon \varepsilon^{\frac{1}{\kappa}}\le|\xi|\le\frac{\delta}{2}\right\}\\ &\ge \liminf_{\varepsilon\downarrow 0}\varepsilon^{\rho} m\left\{\xi\colon \varepsilon\le|\xi|\le\frac{\delta}{2}\right\}>0.\end{aligned}$$

Consequently, by Lemma 7 we have the following result:

**Corollary 1.** *For any $q\ge 1$ there is a constant $C=C(q,T)$ such that*

$$\mathbb{E}\,J_{\phi_\delta}(t)^{-q}\le Ct^{-\frac{\kappa q}{\rho}},\qquad t\in(0,T].$$

*Moreover,*

$$\mathbb{E}\,J_{\phi_\delta}(t)=t\int_{\mathbb{R}}\phi_\delta(\xi)m(\mathrm{d}\xi)<+\infty.$$

## 8. Sharp estimates in the cylindrical $\alpha$-stable case

Here we are concerned with rather general perturbation of $\alpha$-stable case. Indeed in such case we can improve the estimate on $Y(t)$ given in Sect. 6.1. This estimate according to Remark 4 leads to the sharp gradient estimates (6).

Below in (54) we will strengthen hypotheses (8) and (9). In Remark 7 we clarify the validity of the new assumptions in the relevant cylindrical $\alpha$-stable case.

**Lemma 8.** *Let $\alpha\in(0,2)$. Suppose that all the assumptions of Theorem 1 hold with $\rho=\alpha$ and for some $\kappa>1+\alpha/2$. Moreover, suppose that, for the same $\kappa$,*

$$\limsup_{u\to 0^+} u^{-2\kappa+2+\alpha}\int_{|\xi|\le u}\left[|\xi|^{2\kappa}\left(\frac{\rho_j'(\xi)}{\rho_j(\xi)}\right)^2+|\xi|^{2\kappa-2}\right]\rho_j(\xi)\mathrm{d}\xi<+\infty,\tag{54}$$

*and there exists $p\in(1,2)$ such that*

$$\limsup_{u\to 0^+} u^{-p\kappa+p+\alpha}\int_{u\le|\xi|\le r}\left[|\xi|^{p\kappa}\left|\frac{\rho_j'(\xi)}{\rho_j(\xi)}\right|^p+|\xi|^{p\kappa-p}\right]\rho_j(\xi)\mathrm{d}\xi<+\infty.\tag{55}$$

*Then the following estimate holds for the $\mathbb{R}^d$-valued process $Y$ (cf. (43)) :*

$$\mathbb{E}\,|Y(t)|\le C_T\,t^{-\frac{1}{\alpha}},\qquad t\in(0,T].\tag{56}$$

**Remark 7.** We provide a sufficient condition such that all the hypotheses of Lemma 8 hold. To this purpose recall that $\rho_j$ is the $C^1$-density of the Lévy measure $m_j$ associated to the process $Z_j$; such density exists on $(-r,r)\backslash\{0\}$, $r>0$ as in (iii) of Theorem 1.

Moreover, $l_\alpha(\xi):=|\xi|^{-1-\alpha}$ denotes the density of the Lévy measure of a symmetric one-dimensional $\alpha$-stable process, $\alpha\in(0,2)$.

Assume that there is a positive constant $c$ such that, for $\xi \in (-r,r)\setminus\{0\}$,

$$\left|\frac{\rho_j'(\xi)}{\rho_j(\xi)}\right| \le c\left(|\xi|^{-1}+1\right) \qquad \text{and} \qquad c^{-1}l_\alpha(\xi) \le \rho_j(\xi) \le c l_\alpha(\xi), \tag{57}$$

$j=1,\dots,d$. It is easy to check that (57) implies all the assumptions of Lemma 8 with arbitrary $\kappa \in (1+\alpha/2, 1+\alpha)$. Thus under condition (57) we obtain (56) and the sharp gradient estimates (6).

*Proof.* To prove the result we can assume $d=1$ so that $Y_1 = Y$; $\Pi$ is the associated Poisson random measure and we set $m_1 = \mu$ for the corresponding Lévy measure having $C^1$-*density* $\rho_1 = \rho$ on $(-r,r)$.

It is enough to show (56) for small $t$, say $0 < t^{1/\alpha} \vee t \le \delta/2 \le 1$, where $\delta \le r$ is small enough.

Note that $\phi_\delta(\xi) = |\xi|^\kappa$ for $|\xi| \le \delta/2$. Moreover, recall that $\psi_\delta(t) = 1$ for $t \le \delta/2$. Let us fix $\kappa = 1 + \frac{3}{4}\alpha$. We have

$$Y(t) = \frac{D^*\mathbf{1}(t)}{Z^V(t)} - D\frac{1}{Z^V(t)} = \frac{D^*\mathbf{1}(t)}{Z^V(t)} + \frac{DZ^V}{(Z^V(t))^2}.$$

We have

$$\begin{aligned} D^*\mathbf{1}(t) &= -\int_0^t \int_{(-\delta,\delta)} \frac{\phi_\delta'(\xi)\rho(\xi) + \phi_\delta(\xi)\rho'(\xi)}{\rho(\xi)} \widehat{\Pi}(\mathrm{d}s,\mathrm{d}\xi), \\ DZ^V(t) &= \int_0^t \int_{(-\delta,\delta)} \phi_\delta(\xi)\phi_\delta'(\xi)\Pi(\mathrm{d}s,\mathrm{d}\xi), \\ Z^V(t) &= \int_0^t \int_{(-\delta,\delta)} \phi_\delta(\xi)\Pi(\mathrm{d}s,\mathrm{d}\xi). \end{aligned}$$

We are showing that

$$\mathbb{E}\left|\frac{D^*\mathbf{1}(t)}{Z^V(t)}\right| \le C_2 t^{-\frac{1}{\alpha}}. \tag{58}$$

We concentrate on $D^*\mathbf{1}(t)$:

$$D^*\mathbf{1}(t) = I_1(t) + I_2(t), \quad I_1(t) = -\int_0^t \int_{\{t^{1/\alpha}<|\xi|<\delta\}} \frac{\phi_\delta'(\xi)\rho(\xi)+\phi_\delta(\xi)\rho'(\xi)}{\rho(\xi)} \widehat{\Pi}(\mathrm{d}s,\mathrm{d}\xi),$$

$$I_2(t) \quad = -\int_0^t \int_{\{|\xi|\le t^{1/\alpha}\}} \frac{\phi_\delta'(\xi)\rho(\xi) + \phi_\delta(\xi)\rho'(\xi)}{\rho(\xi)} \widehat{\Pi}(\mathrm{d}s,\mathrm{d}\xi).$$

Concerning $I_1(t)$ we can improve some estimates of Sect. 6.1; using the Hölder inequality (because $\xi$ is separated from 0): for the given $p \in (1,2)$ and $q$: $1/p + 1/q = 1$ we have

$$\mathbb{E}\left|I_1(t)\left(Z^V(t)\right)^{-1}\right| \le \left(\mathbb{E}\,|I_1(t)|^p\right)^{1/p}\left(\mathbb{E}\left|Z^V(t)\right|^{-q}\right)^{1/q}.$$

By Corollary 1, there is a constant $C_1$ such that $\mathbb{E}\left|Z^V(t)\right|^{-q} \le C_1 t^{-\frac{\kappa q}{\alpha}}$, recall that $\rho = \alpha$ now. Since $p \in (1,2)$, there exists a positive constant $c$ such that

$$\mathbb{E}\,|I_1(t)|^p \le c\int_0^t \int_{\{t^{1/\alpha}<|\xi|<\delta\}} \left|\frac{\phi_\delta'(\xi)\rho(\xi) + \phi_\delta(\xi)\rho'(\xi)}{\rho(\xi)}\right|^p \rho(\xi)\mathrm{d}\xi\mathrm{d}s,$$

see e.g. Lemma 8.22 in [19]. Since $\phi_\delta(\xi)=|\xi|^\kappa\psi_\delta(\xi)$, it follows that for $|\xi|\le\delta$, $|\phi'_\delta(\xi)|\le C_\delta|\xi|^{k-1}$.

We have by (55)

$$\int_{\{t^{1/\alpha}<|\xi|<\delta\}}\left|\frac{\phi'_\delta(\xi)\rho(\xi)+\phi_\delta(\xi)\rho'(\xi)}{\rho(\xi)}\right|^p\rho(\xi)\mathrm{d}\xi\le C_3(t^{1/\alpha})^{p(\kappa-1)-\alpha}=C_3t^{\frac{p}{\alpha}(\kappa-1)-1}$$

with some constant $C_3$. Combined with the previous inequality, this gives

$$\mathbb{E}\,|I_1(t)|^p\le ct\cdot C_3t^{\frac{p}{\alpha}(\kappa-1)-1}=cC_3t^{\frac{p}{\alpha}(\kappa-1)}.$$

Therefore by the Hölder inequality

$$\mathbb{E}\left|I_1(t)\left(Z^V(t)\right)^{-1}\right|\le C_1^{\frac{1}{q}}t^{-\frac{\kappa}{\alpha}}\cdot(cC_3)^{\frac{1}{p}}t^{\frac{1}{\alpha}(\kappa-1)}=C_4t^{-\frac{1}{\alpha}}. \tag{59}$$

For $I_2(t)$, we proceed in a similar way. Namely, by the Cauchy inequality, isometry formula, Lemma 6 and using (54) we find

$$\begin{aligned}&\mathbb{E}\,|I_2(t)(Z^V(t))^{-1}|\\&\quad\le\left(\mathbb{E}\,(I_2(t))^2\right)^{\frac{1}{2}}\left(\mathbb{E}\,|Z^V(t)|^{-2}\right)^{\frac{1}{2}}\\&\quad\le C_5\left(\int_0^t\int_{\{|\xi|\le t^{1/\alpha}\}}\left(\frac{\phi'_\delta(\xi)\rho(\xi)+\phi_\delta(\xi)\rho'(\xi)}{\rho(\xi)}\right)^2\rho(\xi)\mathrm{d}\xi\right)^{1/2}t^{-\frac{\kappa}{\alpha}}\\&\quad\le C_6\left(t(t^{1/\alpha})^{2(\kappa-1)-\alpha}\right)^{\frac{1}{2}}t^{-\frac{\kappa}{\alpha}}=C_6t^{-\frac{1}{\alpha}},\end{aligned} \tag{60}$$

which completes the proof of (58). Now we are showing that

$$\mathbb{E}\left|\frac{DZ^V(t)}{\left(Z^V(t)\right)^2}\right|\le C_7t^{-\frac{1}{\alpha}}. \tag{61}$$

To this end note that

$$\begin{aligned}&\left|\int_0^t\int_{\{\delta/2<|\xi|\le\delta\}}\phi_\delta(\xi)\phi'_\delta(\xi)\Pi(\mathrm{d}s,\mathrm{d}\xi)\right|\\&\quad\le\left[\int_0^t\int_{\{\delta/2<|\xi|\le\delta\}}\phi^2_\delta(\xi)\Pi(\mathrm{d}s,\mathrm{d}\xi)\right]^{1/2}\left[\int_0^t\int_{\{\delta/2<|\xi|\le\delta\}}\left(\phi'_\delta(\xi)\right)^2\Pi(\mathrm{d}s,\mathrm{d}\xi)\right]^{1/2}\\&\quad\le\int_0^t\int_{\{\delta/2<|\xi|\le\delta\}}\phi_\delta(\xi)\Pi(\mathrm{d}s,\mathrm{d}\xi)\int_0^t\int_{\{\delta/2<|\xi|\le\delta\}}\left|\phi'_\delta(\xi)\right|\Pi(\mathrm{d}s,\mathrm{d}\xi)\\&\quad\le Z^V(t)\int_0^t\int_{\{\delta/2<|\xi|\le\delta\}}\left|\phi'_\delta(\xi)\right|\Pi(\mathrm{d}s,\mathrm{d}\xi).\end{aligned}$$

Hence, as the arguments from the derivation of (59) (recall that for $|\xi| \le \delta$, $|\phi_\delta'(\xi)| \le C_\delta |\xi|^{k-1}$) we obtain

$$\mathbb{E}\left|\frac{\int_0^t \int_{\{\delta/2<|\xi|<\delta\}} \phi_\delta(\xi)\phi_\delta'(\xi)\Pi(\mathrm{d}s,\mathrm{d}\xi)}{\left(Z^V(t)\right)^2}\right| \le \mathbb{E}\,\frac{\int_0^t \int_{\{t^{1/\alpha}<|\xi|<\delta\}} |\phi_\delta'(\xi)|\,\Pi(\mathrm{d}s,\mathrm{d}\xi)}{Z^V(t)}$$
$$\le cC_3C\, t^{\frac{1}{\alpha}(\kappa-1)} t^{-\frac{\kappa}{\alpha}} = C_9 t^{-\frac{1}{\alpha}}.$$

Set

$$K(t) := (Z^V(t))^{-2} \int_0^t \int_{\{\delta/2\ge|\xi|>t^{1/\alpha}\}} \phi_\delta'(\xi)\phi_\delta(\xi)\,\Pi(\mathrm{d}s,\mathrm{d}\xi)$$

and

$$H(t) := (Z^V(t))^{-2} \int_0^t \int_{\{|\xi|\le t^{1/\alpha}\}} \phi_\delta'(\xi)\phi_\delta(\xi)\,\Pi(\mathrm{d}s,\mathrm{d}\xi).$$

Since $\phi_\delta(\xi) = |\xi|^\kappa$ if $|\xi| \le \delta/2$, we have

$$\begin{aligned}
|K(t)| &\le \kappa \left(Z^V(t)\right)^{-2} \int_0^t \int_{\{\delta/2\ge|\xi|>t^{1/\alpha}\}} \phi_\delta^2(\xi)|\xi|^{-1}\,\Pi(\mathrm{d}s,\mathrm{d}\xi)\\
&\le \kappa \left(Z^V(t)\right)^{-2} t^{-\frac{1}{\alpha}} \int_0^t \int_{\mathbb{R}} \phi_\delta^2(\xi)\,\Pi(\mathrm{d}s,\mathrm{d}\xi)\\
&= \kappa \left(Z^V(t)\right)^{-2} t^{-\frac{1}{\alpha}} \left[\left(\int_0^t \int_{\mathbb{R}} \phi_\delta^2(\xi)\,\Pi(\mathrm{d}s,\mathrm{d}\xi)\right)^{1/2}\right]^2\\
&\le \kappa \left(Z^V(t)\right)^{-2} t^{-\frac{1}{\alpha}} \left[\int_0^t \int_{\mathbb{R}} \phi_\delta(\xi)\,\Pi(\mathrm{d}s,\mathrm{d}\xi)\right]^2 = \kappa\, t^{-\frac{1}{\alpha}}.
\end{aligned}$$

We are dealing now with $H(t)$. Since

$$\begin{aligned}
&\int_0^t \int_{\{|\xi|\le t^{1/\alpha}\}} |\phi_\delta'(\xi)\phi_\delta(\xi)|\Pi(\mathrm{d}s,\mathrm{d}\xi)\\
&\le \left(\int_0^t \int_{\{|\xi|\le t^{1/\alpha}\}} (\phi_\delta'(\xi))^2\Pi(\mathrm{d}s,\mathrm{d}\xi)\right)^{1/2} \left(\int_0^t \int_{\mathbb{R}} \phi_\delta^2(\xi)\Pi(\mathrm{d}s,\mathrm{d}\xi)\right)^{1/2}\\
&\le Z^V(t) \int_0^t \int_{\{|\xi|\le t^{1/\alpha}\}} |\phi_\delta'(\xi)|\,\Pi(\mathrm{d}s,\mathrm{d}\xi) = \kappa Z^V(t) \int_0^t \int_{\{|\xi|\le t^{1/\alpha}\}} |\xi|^{\kappa-1}\,\Pi(\mathrm{d}s,\mathrm{d}\xi),
\end{aligned}$$

we have, arguing as in (60), using again (54),

$$\mathbb{E}\,|H(t)| \le \kappa\,\mathbb{E}\,\frac{\int_0^t \int_{\{|\xi|\le t^{1/\alpha}\}} |\xi|^{\kappa-1}\,\Pi(\mathrm{d}s,\mathrm{d}\xi)}{Z^V(t)} \le C_{10} t^{-\frac{1}{\alpha}},$$

which finishes the proof of (61). □

## Acknowledgements

We would like to thank Prof. Jerzy Zabczyk for very useful discussions on the topic. We also would like to thank the anonymous referee for his excellent work pointing out several useful comments and corrections on the previous version of the paper.

**Funding Information** Open access funding provided by Università degli Studi di Pavia within the CRUI-CARE Agreement.



## References

[1] Applebaum, D.: Lévy Processes and Stochastic Calculus, 2nd edn. Cambridge Studies in Advanced Mathematics, Cambridge (2011)

[2] Bally, V., Clement, E.: Integration by parts formula and applications to equations with jumps. Prob. Theo. Rel. Fields **151**, 613–657 (2011)

[3] Bass, R.F., Cranston, M.: The Malliavin calculus for pure jump process and applications to local time. Ann. Probab. **14**, 490–532 (1986)

[4] Bessaih, H., Hausenblas, E., Razafimandimby, P.A.: Ergodicity of stochastic shell models driven by pure jump noise. SIAM J. Math. Anal. **48**, 1423–1458 (2014)

[5] Bismut, J.M.: Calcul des variations stochastique et processus de sauts. Z. Wahrsch. Verw. Gebiete **63**, 147–235 (1983)

[6] Davis, M., Johansson, M.: Malliavin Monte Carlo Greeks for jump diffusions. Stochastic Process. Appl. **116**, 101–1029 (2006)

[7] Dong, Z., Song, Y., Xie, Y.: Derivative formula and coupling property for linear SDEs driven by Lévy processes. Acta Math. Appl. Sin. Engl. Ser. **35**, 708–721 (2019)

[8] Dong, Z., Peng, X., Song, Y., Zhang, X.: Strong Feller properties for degenerate SDEs with jumps. Ann. Inst. Henri Poincaré Probab. Stat. **52**, 888–897 (2016)

[9] Du, K., Zhang, X.: Optimal gradient estimates of heat kernels of stable-like operators. Proc. Am. Math. Soc. **147**, 3559–3565 (2019)

[10] Elworthy, K.D., Li, X.-M.: Formulae for the derivatives of heat semigroups. J. Funct. Anal. **125**, 252–286 (1994)

[11] Fournie, E., Lasry, J.M., Lebuchoux, J., Lions, P.L., Touzi, N.: Applications of Malliavin calculus to monte Carlo methods in finance. Finance Stoch. **3**, 391–412 (1999)

[12] Ishikawa, Y.: Stochastic Calculus of Variations for Jump Processes, De Gruyter Studies in Mathematics, 54, Walter de Gruyter, 2nd ed. (2016)

[13] Ivanenko, D.O., Kulik, A.M.: Malliavin calculus approach to statistical inference for Lévy driven SDE's. Methodolo. Comput. Appl. Prob. **17**(1), 107–123 (2013)

[14] Kawai, R., Takeuchi, A.: Greeks formulas for an asset price model with gamma processes. Math. Finance **21**, 723–742 (2011)

[15] Kulczycki, T., Ryznar, M.: Semigroup properties of solutions of SDEs driven by Lévy processes with independent coordinates, preprint arXiv:1906.07173

[16] Léandre, R.: Régularité de processus de sauts dégénéré. Ann. Inst. H. Poincaré Probab. Statist. **21**, 125–146 (1985)

[17] Norris, J.R.: Integration by parts for jump processes, Séminaire de Probabilité XXII, pp. 271–315, Lecture Notes in Math. 1321, Springer (1988)

[18] Peszat, S., Zabczyk, J.: Strong Feller property and irreducibility for diffusions on Hilbert spaces. Ann. Probab. **23**, 157–172 (1995)

[19] Peszat, S., Zabczyk, J.: Stochastic Partial Differential Equations with Lévy Noise. Cambridge University Press, Cambridge (2007)

[20] Priola, E., Zabczyk, J.: Liouville theorems for nonlocal operators. J. Func. Anal. **216**, 455–490 (2004)

[21] Takeuchi, A.: Bismut-Elworthy-Li-type formulae for stochastic differential equations with jumps. J. Teoret. Probab. **23**, 576–604 (2010)

[22] Wang, F.Y., Xu, L., Zhang, X.: Gradient estimates for SDEs driven by multiplicative Lévy noise. J. Funct. Anal. **269**, 3195–3219 (2015)

[23] Zhang, X.: Derivative formulas and gradient estimates for SDEs driven by $\alpha$-stable processes. Stochastic Process. Appl. **123**, 1213–1228 (2013)

Alexei M. Kulik
Faculty of Pure and Applied Mathematics
Wroclaw University of Science and Technology
Wybrzeże Wyspiańskiego Str. 27
50–370 Wrocław
Poland
e-mail: oleksii.kulyk@pwr.edu.pl

Szymon Peszat
Institute of Mathematics, Jagiellonian University
Łojasiewicza 6
30–348 Kraków
Poland
e-mail: napeszat@cyf-kr.edu.pl

Enrico Priola
Dipartimento di Matematica "F. Casorati", University of Pavia
Via Ferrata, 5
27100 Pavia
Italy
e-mail: enrico.priola@unipv.it

Received: 26 February 2022.
Accepted: 25 September 2022.